The Accuracy of Age Estimation Using Transition Analysis in the Hamann-Todd Collection

Honors Undergraduate Research Thesis

Presented in partial fulfillment of the requirements for graduation *with honors research distinction* in Anthropological Sciences at The Ohio State University

by

Allyson M. Simon

The Ohio State University

April 2020

Advisor: Dr. Mark Hubbe, Department of Anthropology

The Accuracy of Age Estimation Using Transition Analysis in the Hamann-Todd Collection

Allyson M. Simon

Thesis Defense Committee

Mark Hubbe, Ph. D.

Department of Anthropology

Debra Guatelli-Steinberg, Ph. D.

Department of Anthropology

Christa B. Teston, Ph. D.

Department of English

Table of Contents

I. Abstract

Trying to reconstruct the demographic parameters of past populations using skeletal data is challenging for several reasons. First, skeletal traits are proxy to biological age, which does not always reflect chronological age. Second, different parts of the skeleton age at different rates, making age estimation problematic when based on methods that only incorporate a few anatomical features. A proposed solution to address this limitation is transition analysis (TA), a multifactorial method of age estimation that incorporates information from the pubic symphysis, iliac auricular area, and cranial sutures. However, despite its methodological refinement, TA has been shown to have varying degrees of accuracy when applied to different known-age skeletal samples. This thesis contributes to the discussion about TA's accuracy by estimating the age for 221 individuals from the Hamann-Todd Collection. We contrasted the maximum likelihood estimates generated through TA to the known ages, and analyzed the absolute error for the entire sample, as well as according to sex and ancestry. Estimates show an average absolute error of 11.6 (SD=10.3) years, with white individuals' average absolute errors (14.1 years) being significantly higher than black individuals' (9.1 years; Independent Samples T-Test $p<0.001$). There are no significant differences in the absolute errors between sexes (males = 11.2; females = 12.1; Independent Samples T-Test $p=0.526$). A weak to moderate positive correlation was found between known age and absolute error for white males ($R^2=0.3688$; $p<0.001$), white females ($R^2=0.2590$; $p<0.001$), and black males ($R^2=0.1379$; $p=0.006$). The accuracy of transition analysis age estimates varied when each anatomical region was analyzed independently, but the combination of all three anatomical regions yielded the most accurate age estimates. These findings further support that TA accuracy depends on the prior distribution used

and that in the case of the Hamann-Todd Collection, the accuracy for white individuals is more influenced by this limitation than when black individuals are analyzed.

II. Introduction

Accurate and unbiased age estimations are vital for reconstructing demography in past populations. However, there are numerous challenges with current age estimation methods and applications. Namely, the objective of trying to reconstruct the past with skeletal data is complicated by the fact that biological age does not always reflect chronological age. In other words, chronological age is "a proxy" for an individual's biological rate of aging and at any given chronological age, individuals may show various biological ages within a population (Couoh, 2017). The goal of age estimation methods is to reliably correlate biological or skeletal age with chronological age (Couoh, 2017). We are interested in chronological age in order to reconstruct past demography, but we do not have access to chronological age from the skeleton because it is a human invention. Thus, we can use biological aging, which is visible in different anatomical regions of the skeleton, to infer chronological age. But, biological age can be accelerated or decelerated based on different factors.

The discrepancy between chronological age and true biological age can be attributed to an array of causes including genetics, physical activity levels, body size, stature, as well as environmental conditions (Buckberry and Chamberlain, 2002; Couoh, 2017; Merritt, 2015; Moraitis et al., 2014). Age estimation methods attempt to infer biological ages, which can be compared with known chronological ages to ascertain "inaccuracy" or age estimation error. However, due to the numerous variables that influence the onset of degenerative stages and the rate of skeletal aging processes, it is difficult to determine whether the age estimation method is poorly capturing biological traits of age or if high variability among individuals and populations is causing the discord between estimated biological age and known chronological age. In sum, the rate of skeletal aging, and the rate of aging for specific anatomical regions of the skeleton, is

highly variable at the individual and population level. This poses a challenge for age estimation methods to be reliable in different populations and individuals with considerable variation in factors that influence the onset and rate of skeletal aging processes.

Adult age estimation is particularly difficult because age-related degenerative skeletal changes are more variable from individual to individual than age-related developmental skeletal stages, and is influenced by numerous variables including, but not limited to, genetics, environment, and level of physical activity (Boldsen et al., 2002; Buckberry, 2015; Bullock et al., 2013). Moreover, the rate of bone remodeling varies considerably between sexes and among different populations, which could affect age estimation (Gocha et al., 2019). Furthermore, different parts of the skeleton age at different rates due to varying physical activity patterns, sex, and ancestry. This makes age estimation particularly difficult when the established methods only incorporate few anatomical features. Traditional age estimation methods have no standardized method for combining various anatomical regions of the skeleton (Bullock et al., 2013). Furthermore, traditional age estimation methods tend to underestimate the ages of older individuals and rely on wide intervals (Cappella et al., 2017; Milner and Boldsen, 2012). It has been shown that multifactorial methods of age estimation, such as transition analysis (TA), tend to increase accuracy and control variation among the different stages of aging at different anatomical locations within an individual skeleton (Franklin, 2010). However, limited studies have been conducted to investigate the accuracy of TA and other multifactorial methods (Clark et al., 2019).

TA is a multifactorial method of age estimation that uses the pubic symphysis, the iliac auricular area, and cranial sutures (Boldsen et al., 2002; Milner and Boldsen, 2011). The method was designed using the Terry and Coimbra skeletal collections to define the various component

stages of aging at each of the three skeletal features examined. The ABDOU computer program calculates estimated ages (maximum likelihood point estimates) and a 95% confidence interval (Boldsen et al., 2002; Milner and Boldsen, 2012). The standards for TA account for both archaeological and forensic populations. TA does not require observations of all skeletal traits, and can therefore be used on partial skeletons, a common reality in both archaeological and forensic settings. Although, reduced number of traits available for scoring may affect accuracy of the estimation (Jooste et al., 2016). However, TA has been shown perform well without all traits being present (Godde and Hens 2012; Hens and Godde, 2016; Konigsberg, 2008).

TA also addresses issues of age mimicry, a problem initially described by Bocquet-Appel and Masset (1982), in which age estimates of the target population are biased by the sample used to develop the methods for age estimation (Boldsen et al. 2002; Bullock et al., 2013; Godde and Hens, 2012). Age mimicry is a challenge to age estimation because it is an inherent form of bias that impedes our ability to reconstruct accurate mortality patterns in various populations (Boldsen et al., 2002). Regression analysis, which most traditional age estimation methods are based on, biases the mortality patterns of the target population to that of the reference population (Bolden et al., 2002). Conversely, TA employs Bayesian modeling, which aims to reduce the bias associated with age mimicry. TA calculates the probability of transitioning from one stage of skeletal aging to another, hence the name *transition* analysis. The use of Bayesian modeling has been found to increase the accuracy of traditional aging methods, as well as of TA (Godde and Hens, 2012; Hens and Godde, 2016). Yet, sometimes age mimicry does not significantly influence age-at-death distributions (Clark et al., 2019; Moraitis et al., 2014).

Many studies have demonstrated the increased accuracy of TA compared to traditional aging methods (Bullock et al., 2013; Jooste et al., 2016). However, many studies have also

demonstrated TA to be relatively inaccurate and contend that due to variation among populations, the informative prior distributions are inappropriate for different target samples (Jooste et al., 2016; Milner and Boldsen, 2011; Milner and Boldsen, 2012; Xanthopoulou et al., 2018). For example, Xanthopoulou et al. (2018) concluded that TA was less accurate in a contemporary Greek skeletal assemblage compared to traditional aging methods. One of the cited reasons for these results is inter and intra population differences which make TA more suitable for age estimation in some populations than in others. However, although it has been observed that variation among populations in the onset and rate of skeletal aging processes can affect the accuracy of TA (and other age estimation methods; Moraitis et al., 2014), researchers have generally concluded that the effect is not significant (Franklin, 2010; Garvin et al., 2012; Konigsberg et al., 2008).

Other studies claim that while TA increases the accuracy of age estimation, the method lacks precision, especially when few anatomical traits are scored (Jooste et al., 2016; Milner and Boldsen, 2012). On the other hand, Milner and Boldsen's (2012) validation study of TA found that the method is better at reconstructing past demography than individual age-at-death estimations. Thus, TA is better equipped to illustrate population trends in mortality, but TA is still lacking in its ability to obtain accurate and precise biological profiles for individual skeletons. Together, these studies illustrate the lack of consensus that still exists in the literature regarding the efficacy and accuracy of TA to estimate age in skeletal remains.

This thesis aims to contribute to this discussion by investigating whether the standards established for transition analysis, the informative prior distributions, are valid for the Hamann-Todd Collection. This thesis assesses the absolute age estimation error of transition analysis for the Hamann-Todd Collection, focusing on whether the established standards for transition

analysis for ancestry and sex are accurate in another population. It is expected that the standards for transition analysis will reflect age-at-death in the Hamann-Todd Collection.

III. Materials & Methods

The Hamann-Todd Osteological Collection is housed in the Cleveland Museum of Natural History, formed from 1912 to 1938, and contains over 3,000 human skeletons (Meindl et al., 1983). A majority of the individuals that compose the Hamann-Todd Collection were from lower socio-economic backgrounds and many worked as unskilled laborers (de la Cova, 2010). These conditions likely contributed to the significant rate of trauma and pathology in the collection (de la Cova, 2011). Their remains were often added to the collection as unclaimed bodies from the county morgue or municipal hospitals from the Cleveland area (de la Cova, 2010). Most of the individuals in the Hamann-Todd Collection have documentation of age at death, biological sex, race, and cause of death. However, some stated ages may be approximations since a few cadavers lacked reliable medical records and documentation (Meindl et al., 1983). While the Hamann-Todd Collection is representative of both sexes, all ages, and various racial backgrounds, it is largely composed of adult white males.

The sample was comprised of 56 black females, 56 white females, 54 black males, and 55 white males. These samples were generated randomly from the full list of complete skeletons in the Hamann-Todd Collection, and stratified for sex and ancestry[1]. Although TA is designed so that partial skeletons can be used, this sample aimed to use only complete skeletons, since past

[1] The use of the term "ancestry" here is actually referring to race, which was assigned to the skeletons in this collection from medical and autopsy documentation as either "black" or "white." However, the true *ancestries* of the individuals in this collection are unknown. Race is cautiously used as a proxy for ancestry in this research.

research has found decreased accuracy of the method with less available traits to score (Jooste et al., 2016; Milner and Boldsen, 2012). The age distribution of individuals in each subgroup (black females, white females, black males, and white males), shown in Table 1 and Figure 1, reflects the actual demographic make-up of the Hamann-Todd Collection. For example, the greater proportion of white females in the 70+ age category mirrors the demographic make-up of the Hamann Todd Collection, as does the lower proportion of black females after approximately the 50-54 age category (Figure 1).

Table 1. Summary of Sample Demographics

Age Category	Black Females	White Females	Black Males	White Males
20-24	7	1	5	1
25-29	10	3	2	0
30-34	6	3	7	5
35-39	5	6	6	5
40-44	7	5	11	8
45-49	7	4	8	10
50-54	4	8	3	9
55-59	1	3	4	3
60-64	3	3	3	5
65-69	2	2	3	5
70+	4	18	2	4
Total	56	56	54	55

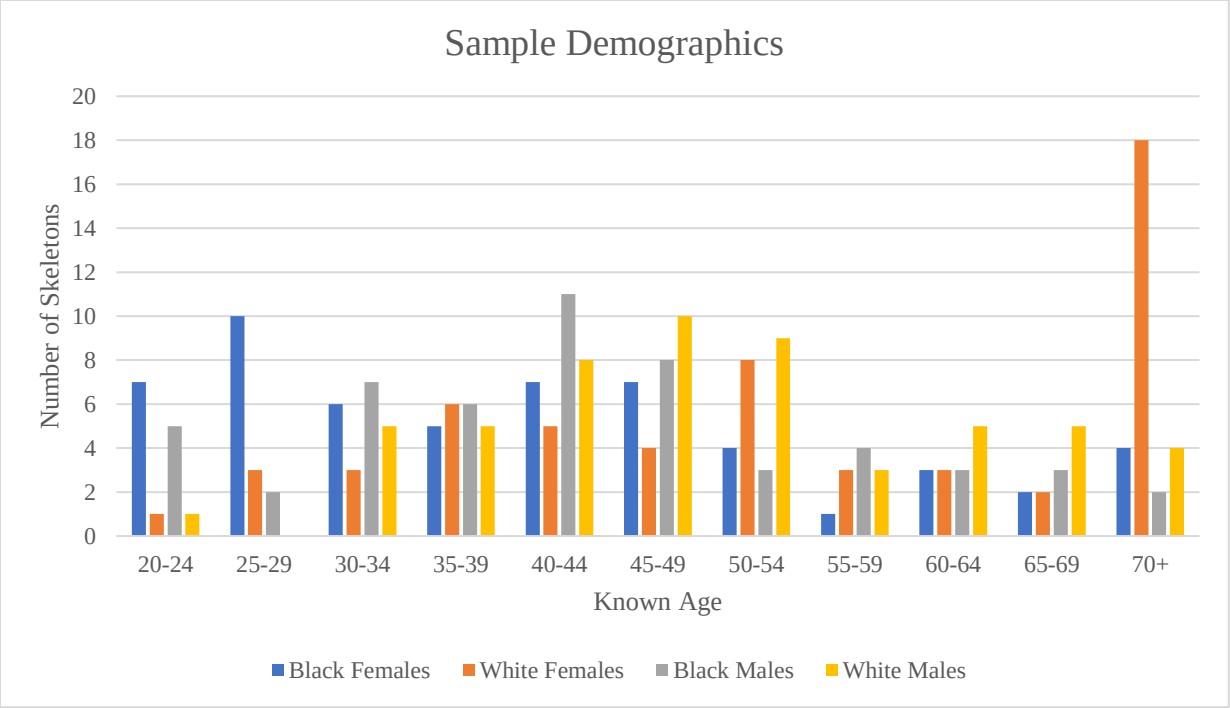


Figure 1. Summary of the demographic information of the sample included in this study.

Age at death was estimated for all individuals in the sample using TA. The procedures outlined in Boldsen et al. (2002) were used. TA age estimation is based on the scoring of 19 variables from three anatomical regions. There are five variables scored from the cranial sutures, five variables scored from the pubic symphysis, and nine variables scored from the auricular surface. Bilateral traits were scored when applicable, bringing the total number of possible scores from one skeleton to 36. Scores were inputted into the ADBOU software available at http://math.mercyhurst.edu/~sousley/Software/.

Intra-observer error was calculated for each trait scored using Cohen's Kappa values (Table 2). Scores were highly consistent between the observer's first and second scoring, with the lowest Kappa value obtained for superior surface morphology (0.643). It should be noted that Fojas et al. (2018) similarly found that the superior surface morphology was the trait least

consistently scored between observers. Cranial suture scores were the most consistently scored. The pubic symphysis and auricular surface scores were less repeatable but still very consistent between the first and second scoring.

Table 2. Summary of Kappa Values for Intra-Observer Error

Trait	Kappa Value
Coronal Pterica	1.000
Sagittal Obelica	1.000
Lambdoidal Asterica	1.000
Interpalatine	1.000
Zygomaticomaxillary	1.000
Symphyseal Relief	0.833
Symphyseal Texture	1.000
Superior Apex	0.853
Ventral Symphyseal Margin	0.839
Dorsal Symphyseal Margin	1.000
Superior Demiface Topography	1.000
Inferior Demiface Topography	1.000
Superior Surface Morphology	0.643
Middle Surface Morphology	0.855
Inferior Surface Morphology	0.841
Inferior Surface Texture	1.000
Superior Posterior Illiac Exostoses	0.825
Inferior Posterior Illiac Exostoses	1.000
Posterior Exostoses	1.000

Accuracy of an estimate was determined by whether the known age fell within the 95% confidence interval generated by the ADBOU program using the appropriate informative prior distribution. The known ages were then compared with maximum likelihood point estimates. The absolute error (in years) for a given age estimate was calculated by taking the absolute value of the difference between known age and the computed maximum likelihood. A T-test was performed on the sub-sample absolute error means to test for statistically significant differences between subgroups (i.e. black females, white females, black males, and white males). These

comparisons were made to investigate whether the differences in mean absolute error between the sexes or between white and black individuals are statistically significant. Statistically significant differences would suggest that TA is more accurate in one population compared to another. Correlation tests were used to test the association between increasing known age and absolute error, to determine whether TA is less accurate in older individuals.

Finally, to test the accuracy of each anatomical region (i.e. cranial sutures, pubic symphysis, and auricular surface), each anatomical region was analyzed independently. Since TA is designed to facilitate age estimation on partial skeletons, the accuracy of each anatomical region alone is important. This was examined by looking at the absolute error associated with each individual anatomical region when used to estimate age without scores from the other regions (as if those other traits were not present).

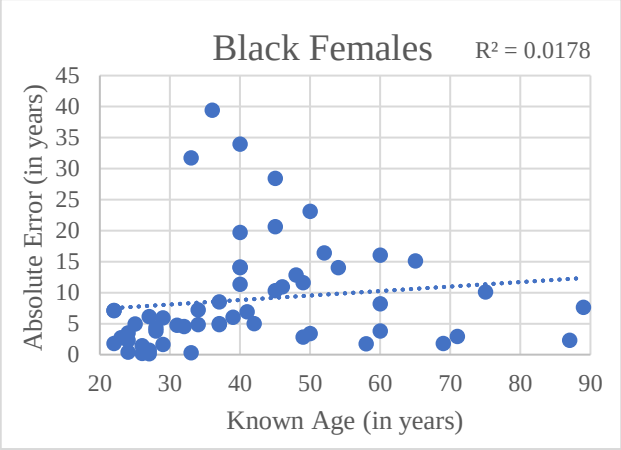
IV. Results

Overall, transition analysis did not perform well in this sample. Differences between maximum likelihood estimates and known ages show a mean absolute error of 11.6 years (SD=10.3), with the largest absolute error being 48 years. Estimates were considered "accurate" if the known age fell within the 95% confidence interval generated by the ADBOU program. Of the entire sample (n=221), 59 (26.7%) individuals' known ages fell outside the estimated 95% confidence interval. Ten incomplete skeletons (damaged or missing bones) were included in these analyses, of which six have known ages outside of the 95% confidence interval of TA. While this suggests that incomplete skeletons have a high probability of having their age poorly estimated by TA, their inclusion in the analyses did not significantly influenced results because the mean absolute error when all partial or damaged skeletons are excluded from the sample is very similar to the observed for the whole sample (11.1 years, SD=9.8).

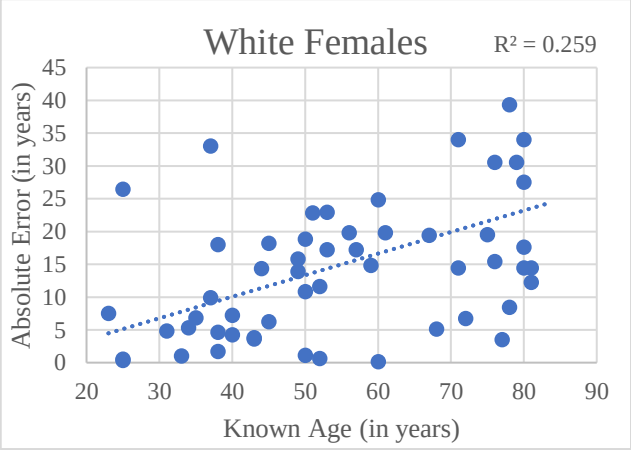
A weak to moderate positive correlation was found between known age of the individual and absolute error observed from the TA maximum likelihood estimate for three of the four subgroups analyzed (Figures 2a-d; Table 3). Increasing age explains very little of the variation in TA error for black females (R^2=0.018; Rho= +0.133; p=0.33). However, the correlation between increasing age and absolute error was found to be statistically significant using a Spearman's Correlation Test for black males (R^2=0.138; Rho= +0.373; p=0.006), white females (R^2=0.259; Rho= +0.509; p<0.001), and white males (R^2=0.369; Rho= +0.607; p<0.001), suggesting that increasing age plays a greater role in inaccurate estimates for black males and white individuals.

Table 3. Spearman's Correlation Test Values

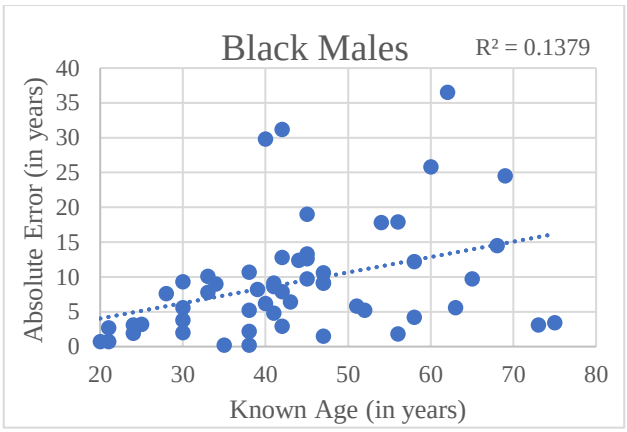
Subgroup	R^2 Value	Rho Value	P-value
Black females	0.018	+0.133	0.327
White females	0.259	+0.509	<0.001
Black males	0.138	+0.371	0.006
White males	0.369	+0.607	<0.001



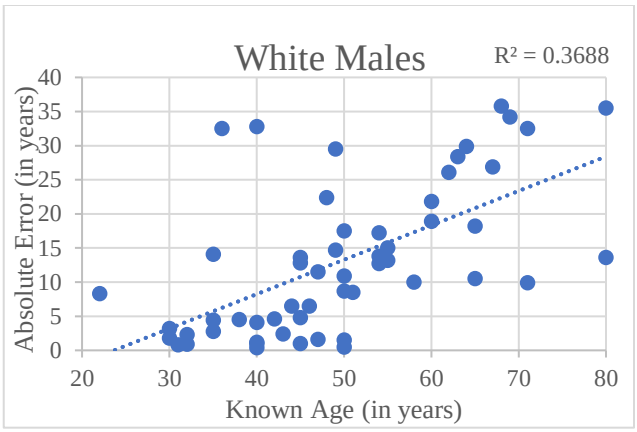
(a)



(b)



(c)



(d)

Figure 2(a-d). Correlations between known age (in years) and absolute error (in years) for black females (a), white females (b), black males (c), and white males (d)

The accuracy of TA age estimates using each anatomical region individually varied considerably (Figure 3). Cranial sutures were by far the least accurate and least precise anatomical region, having an average of absolute error of 23.15 years. Cranial sutures tended to overestimate the ages of individuals (most errors are positive): even when not using absolute values, the average of error is still 18.8 years (Figure 4). The pubic symphysis was the most accurate of the three traits when analyzed alone. However, it still had an average absolute error of 15.3 years, with a trend of underestimating age (average of -13.6 years for raw observed error) (Figure 4). The auricular surface had an average absolute error of 18.5 years and overestimated age by an average of 11.7 years when raw data was analyzed (Figure 4).

An analysis of the absolute error when only the pubic symphysis and auricular surface data are used reveals a higher mean absolute error of 12.3 years. Thus, the combination of only the pubic symphysis and auricular surface do not improve the accuracy of TA estimates in this sample. Overall, there was a substantial amount of variation in the accuracy between each anatomical region, but there is not one anatomical region that was accurate enough on its own to

be used to the exclusion of the other anatomical regions. In other words, findings suggest that it is not more or equally as accurate to use only the pubic symphysis or only the auricular surface with TA, i.e., age estimates that employed all anatomical traits had smaller errors on average. Thus, the combination of cranial sutures, pubic symphysis, and auricular surface improves age estimation in this sample.

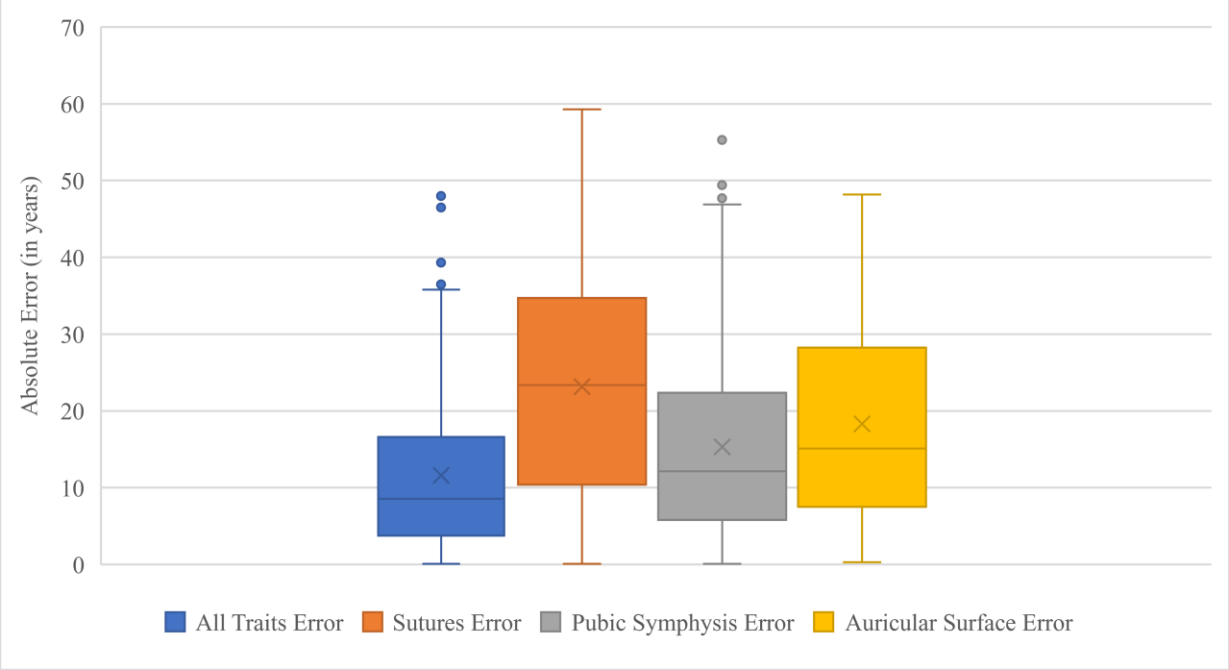


Figure 3. Box-plot showing the variation in accuracy for each anatomical region

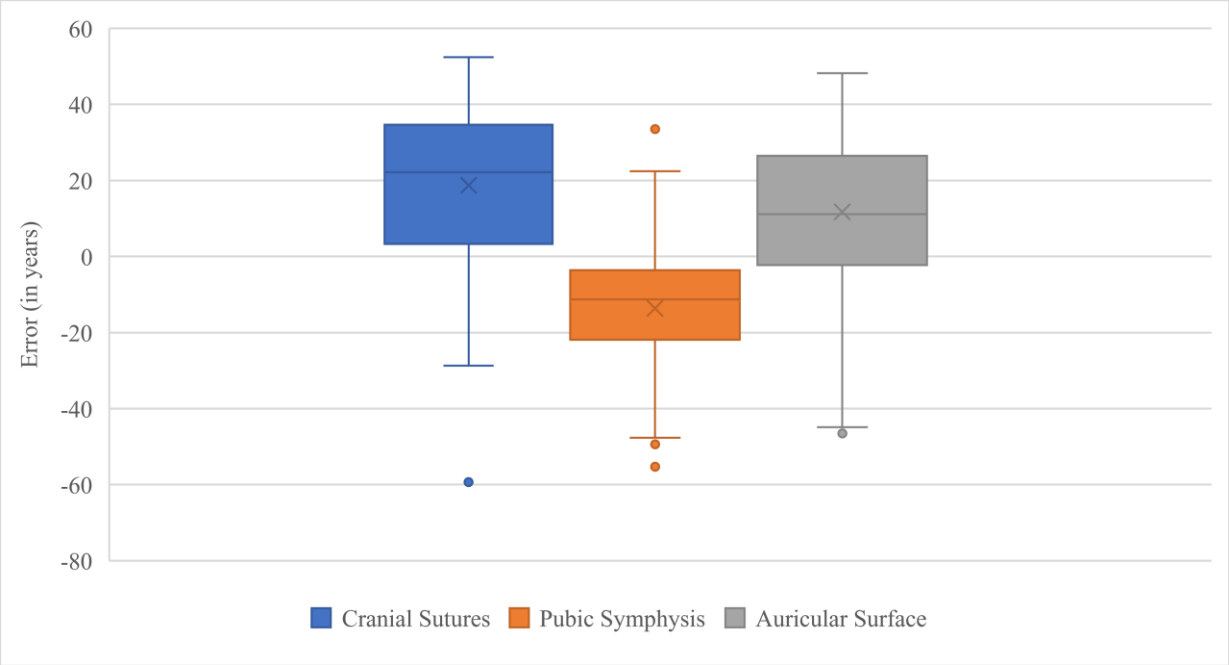


Figure 4. Box-plot showing the error among the different anatomical regions analyzed. Error is the difference (in years) between the maximum likelihood estimates and known age.

Accuracy of TA varied among the four subgroups. In particular, TA performed significantly worse on white individuals of both sexes. The results of the T-test show that the average absolute error for white individuals (14.1 years) is significantly higher than for black individuals (average absolute error = 9.1 years; $p<0.001$). This trend is consistent when black females (average absolute error=8.920 years) is compared with white females (average absolute error = 15.189 years; $p=0.002$), and black males (average absolute error=9.244 years) compared with white males (average absolute error = 13.067; $p=0.040$). Maximum likelihood estimates for white individuals tended to underestimate age significantly more than for black individuals. There are no significant differences in the absolute errors between sexes (males = 11.2; females= 12.1; $p=0.526$). Consistently, there are also no significant differences in the average absolute

errors between white females and white males (p=0.327) and between black females and black males (p=0.841). A summary of these p-values can be viewed in Table 4.

Table 4. Summary of T-Tests on Mean Absolute Error between Subgroups

Test	T-value	P-value
Females vs. Males	0.635	0.526
White vs. Black	3.761	<0.001
Black females vs. White females	-3.190	0.002
Black males vs. White males	-2.077	0.040
White females vs. White males	0.985	0.327
Black females vs. Black males	-0.201	0.841

As demonstrated in Figure 5, black females had the least absolute error on average, followed by black males, white males, and white females, respectively. Given the statistical significance of the difference in absolute error between white and black individuals, TA was less accurate for white individuals in this sample. Overall, the results of this study show that TA was not very accurate in the Hamann-Todd Collection, especially with regards to white individuals in the sample.

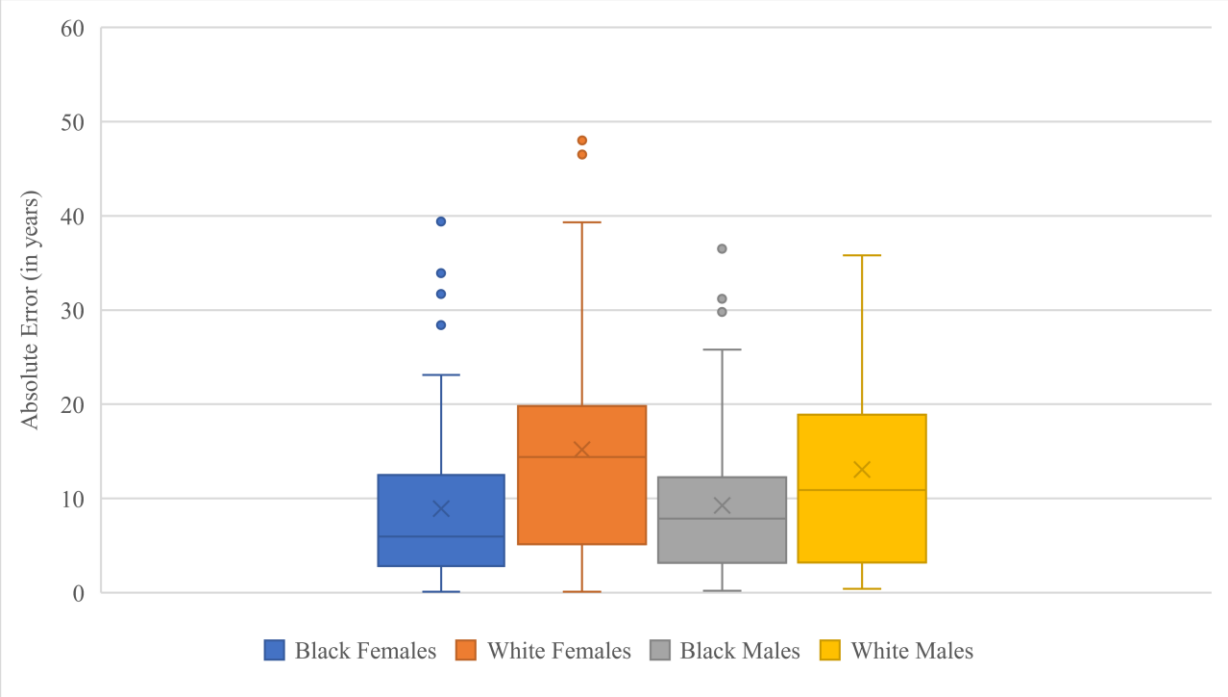


Figure 5. Boxplot showing the absolute error (in years) between observed age and TA estimate for each sub-group.

V. Discussion

The results of this thesis demonstrate that TA age estimates do not accurately reflect known age in the Hamann-Todd Collection. The mean absolute error of TA for this sample is 11.6 years (SD=10.3). In comparison, Jooste et al. (2016) found a mean absolute error of 10.4 years using TA. More importantly, 26.7% of individuals' known ages fell outside the estimated 95% confidence interval.

Furthermore, while previous studies have found the effects of ancestry to be minimal on the accuracy of age estimation (Franklin, 2010; Garvin et al., 2012; Konigsberg et al., 2008), the findings of this thesis disagree with them. Age estimates were less accurate for white individuals

in the sample ($p<0.001$), suggesting that the informative prior distribution was less appropriate for this subgroup.

The results of the analysis of each anatomical region are consistent with previous studies. The use of all three features in TA yields the most accurate age estimates (Lopez-Cerquera and Casallas, 2018; Milner and Boldsen, 2012). When analyzed independently, cranial sutures were the least accurate and least precise. Numerous studies have similarly found that cranial sutures are the least accurate and precise trait (Milner and Boldsen, 2012), with some finding that TA age estimates are more accurate without data from the cranial sutures (Jooste et al., 2016; Xanthopoulou et al., 2018). However, the results from the analysis of absolute error using only the pubic symphysis and auricular surface, excluding cranial sutures, was still found to be less accurate than when all of the anatomical regions were used for this sample.

Among the anatomical regions analyzed individually, the pubic symphysis was the most accurate, although the pubic symphysis did tend to underestimate age-at-death. This is consistent with the findings of Lopez-Cerquera and Casallas (2018) and Milner and Boldsen (2012). The auricular surface was the second most accurate anatomical trait but tended to overestimate age on average. This is also consistent with the findings of Lopez-Cerquera and Casallas (2018). Contrarily, Jooste et al. (2016) found that the auricular surface was the most accurate of the anatomical traits.

These results conflict with the findings of Xanthopoulou et al. (2018), who found that the cranial sutures systematically underestimated age while the pubic symphysis and auricular surface both overestimated age. Overall, there is insufficient literature on the topic to conclusively determine the optimal anatomical regions for age estimation and how the rate at which the various degenerative stages of aging affect accuracy among different populations.

Furthermore, as previously stated, TA age estimates in the Hamann-Todd tended to be the most accurate when all three regions were used. However, Jooste et al. (2016) and Xanthopoulou et al. (2018) advocate for the exclusion of cranial sutures, yielding more accurate age estimates without these scores.

Overall, TA estimates did not accurately reflect known age of individuals in the sample, especially for white individuals. This is likely due to variation in the rate of degenerative processes between individuals and populations. Findings suggest that the limitations of age estimation are not entirely due to the methods being used, but also reflect unknowns about how individuals age and what affects this process. The results of this research imply that TA accuracy depends partially on the informative prior distribution used, which is less accurate for white individuals in this sample from the Hamann-Todd Collection. Thus, the informative prior distribution must be similar to the target sample for the method to shower better accuracy. This is consistent with the findings of Milner and Boldsen (2012) and Godde and Hens (2012).

While TA is better able to address the issues of higher inaccuracy of age estimates among older individuals, it is not immune to the same pitfalls of traditional methods. The weak to moderate positive correlations between known age and error for most subgroups show that advanced age remains a source of error for estimating age, despite being drastically reduced compared to commonly used traditional age estimation methods.

The large degree of variation between anatomical regions when analyzed independently of each other, reinforces the need for multifactorial methods for age estimation. Multifactorial approaches to age estimation are undoubtedly still more accurate than estimates generated from the examination of one feature alone (Buckberry, 2015). Moreover, numerous studies have found

that Bayesian modeling increases the accuracy of age estimates (Godde and Hens, 2012; Hens and Godde, 2016).

Age estimation continues to be one of the most significant challenges to bioarchaeological and paleodemographic research (DeWitte, 2018). Accurate age estimates are imperative in order to reconstruct past demography and generate unbiased data. Furthermore, lack of uniformity within the field of anthropology regarding which age estimation techniques to use is problematic when attempting to compare studies using one method to studies using another method (Buckberry, 2015; Cappella et al., 2017, Clark et al., 2019). This is important because traditional age estimation methods and TA can yield significantly different age-at-death distributions for past populations (Clark et al., 2019).

Nevertheless, developing reliable age estimation methods remains problematic (Couoh, 2017). It should be considered that biological age rarely correlates perfectly with chronological age. Thus, current methods of age estimation should be relied upon with caution. As stated by Couoh (2017), the current practice is to blame the age estimation methods for inadequately or unreliably correlating skeletal age with chronological age despite the diversity of populations to which the methods are being applied. Thus, it may be more constructive to consider inaccurate age estimates as the result of complex factors (e.g. environment, activity levels, genetics, etc.) contributing to the immense variability, on the individual and population level, of the onset and rate of skeletal aging processes, as opposed to the age estimation methods themselves (Moraitis et al., 2014). The influence of these factors on the onset stages and rate of degenerative skeletal aging is still not fully understood, which poses a particular challenge for existing age estimation methods to be equipped to accommodate complexity and variation.

The results of this study suggest that TA is less accurate than previously thought and requires an appropriate informative prior distribution in order to be relatively accurate. TA is especially less accurate for white individuals within the Hamann-Todd Collection. Further studies are needed to assess the accuracy of TA in other populations.

VI. Conclusion

The significant differences observed between the absolute error of TA on white individuals compared to black individuals indicate that TA accuracy depends on the informative prior distribution used. This is important when applying the method to different populations because, as depicted in the results, TA is not similarly accurate among the subgroups studied. TA was significantly less accurate for white individuals in this sample. Therefore, past conclusions that ancestry does not significantly affect the accuracy of age estimation methods is not supported by these findings. This is particularly concerning when applying TA to target samples that are dissimilar to the established informative prior distributions. This finding is of special relevance for the analysis of prehistoric samples, since it is difficult to know whether the informative prior distributions are appropriate for those populations. Consequently, the method inadequately addresses population variation and requires further refinement and further research is needed on the accuracy of TA age estimation. In conclusion, the findings of this thesis reinforce the need for multifactorial, accurate, and precise age estimation methods, but also highlights the unresolved challenges facing age estimation.

References

Bocquet-Appel, J.-P., & Masset, C. (1982). Farewell to Paleodemography. *Journal of Human Evolution, 11*, 321-333.

Boldsen, J., Milner, G., Konigsberg, L. W., & Wood, J. W. (2002). Transition Analysis: A New Method for Estimating Age from Skeletons. In R. D. Hoppa, & J. W. Vaupel, *Paleodemography: Age Distributions from Skeletal Samples.* Cambridge University Press.

Buckberry, J. (2015). The (mis)use of adult age estimates in osteology. *Annals of Human Biology*, *42*(4), 321-329.

Bullock, M., Marquez, L., Hernandez, P., & Ruiz, F. (2013). Paleodemographic Age-at-Death Distributions of Two Mexican Skeletal Collections: A Comparison of Transition Analysis and Traditional Aging Methods. *American Journal of Physical Anthropology, 152*, 67-78.

Cappella, A., Cummaudo, M., Arrigoni, E., Collini, F., & Cattaneo, C. (2017). The Issue of Age Estimation in a Modern Skeletal Population: Are Even the More Modern Current Aging Methods Satisfactory for the Elderly? *Journal of Forensic Sciences, 62*(1), 12-17.

Clark, M. A., Simon, A., & Hubbe, M. (2019). Aging methods and age-at-death distributions: Does Transition Analysis call for a re-examination of bioarchaeological data? *International Journal of Osteoarchaeology*.

Couoh, L. R. (2017). Differences between biological and chronological age-at-death in human skeletal remains: A change of perspective. *American Journal of Physical Anthropology, 163*, 671-695.

de la Cova, C. (2010). Cultural Patterns of Trauma among 19th-Century-Born Males in Cadaver Collections. *American Anthropologist, 112*(4), 589-606.

de la Cova, C. (2011). Race, Health, and Disease in 19th-Century-Born Males. *American Journal of Physical Anthropologists, 144*, 526-537.

DeWitte, S. N. (2018). Demographic anthropology. *American Journal of Physical Anthropology, 165*, 893-903.

Dirkmaat, D. C., Cabo, L. L., Ousley, S. D., & Symes, S. A. (2008). New Perspectives in Forensic Anthropology. *Yearbook of Physical Anthropology, 51*, 33-52.

Falys, C. G., Schutkowski, H., & Weston, D. A. (2006). Auricular Surface Aging: Worse Than Expected? A Test of the Revised Method on a Documented Historical Skeletal Assemblage. *American Journal of Physical Anthropology, 130*, 508-513.

Fojas, C. L., Kim, J., Minksy-Rowland, J. D., & Algee-Hewitt, B. F. (2018). Testing inter-observer reliability of the Transition Analysis aging method on the William M. Bass forensic skeletal collection. *American Journal of Physical Anthropology, 165*, 183-193.

Franklin, D. (2010). Forensic age estimation in human skeletal remains: Current concepts and future directions. *Legal Medicine, 12*, 1-7.

Garvin, H. M., Passalacqua, N. V., Uhl, N. M., Gipson, D. R., Overbury, R. S., & Cabo, L. L. (2012). Developments in Forensic Anthropology: Age-at-Death Estimations. In D. C. Dirkmaat, *A Companion to Forensic Anthropology* (pp. 202-223). Wiley-Blackwell Publishing.

Gocha, T. P., Robling, A. G., & Stout, S. D. (2019). Histomorphometry of Human Cortical Bone: Applications to Age Estimation. In M. A. Katzenberg, & A. L. Grauer, *Biological Anthropology of the Human Skeleton.* John Wiley & Sons, Inc.

Godde, K., & Hens, S. M. (2012). Age-at-Death Estimation in an Italian Historical Sample: A Test of the Suchey-Brooks and Transition Analysis Method. *American Journal of Physical Anthropology, 149*, 259-265.

Hen, S. H., & Godde, K. (2016). Auricular Surface Aging: Comparing Two Methods that Assess Morphological Change in the Ilium with Bayesian Analyses. *Journal of Forensic Sciences, 61*, S30-S38.

Jooste, N., L'Abbe´, E. N., Pretorius, S., & Steyn, M. (2016). Validation of transition analysis as a method of adult age estimation in a modern South African sample. *Forensic Science International, 266*, 580.e1–580.e7.

Konigsberg, L. W., Herrmann, N. P., Wescott, D. J., & Kimmerle, E. H. (2008). Estimation and Evidence in Forensic Anthropology: Age-at-Death. *Journal of Forensic Science, 53*(3), 541-557.

Lopez-Cerquera, M. A., & Casallas, D. A. (2018). Age Estimation in a Colombian Modern Skeletal Sample: A Test of the Transition Analysis Method. *Forensic Anthropology, 2*(1), 22-28.

Meindl, R. S., Lovejoy, C. O., & Mensforth, R. P. (1983). Skeletal Age at Death: Accuracy of Determination and Implications for Human Demography. *Human Biology, 55*(1), 73-87.

Merritt, C. E. (2015). The Influence of Body Size on Adult Skeletal Age Estimation Methods. *American Journal of Physical Anthropology, 156*, 35-57.

Milner, G. R., & Boldsen, J. L. (2011). Transition Analysis Age Estimation: Scoring Manual.

Milner, G. R., & Boldsen, J. L. (2012). Transition Analysis: A Validation Study With Known-Age Modern American Skeletons. *American Journal of Physical Anthropology, 148*, 98-110.

Moraitis, K., Zorba, E., Eliopoulos, C., & Fox, S. C. (2014). A Test of the Revised Auricular Surface Aging Method on a Modern European Population. *Journal of Forensic Sciences, 59*(1), 188-194.

Xanthopoulou, P., Valakos, E., Youlatos, D., & Nikita, E. (2018). Assessing the accuracy of cranial and pelvic aging methods on human skeletal remains from a modern Greek Assemblage. *Forensic Science International, 286*, 266.e1-266.e8.